



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT : Mary Benoit  
SERIAL NO. : 09/494,670 EXAMINER : Shawn S. An  
FILED : January 31, 2000 ART UNIT : 2613  
FOR : DESCRIPTOR FOR A VIDEO SEQUENCE AND IMAGE RETRIEVAL  
SYSTEM USING SAID DESCRIPTOR

AF/2613  
\$  
2200  
#12  
3-16-04  
P.2

APPEAL BRIEF TRANSMITTAL LETTER

**RECEIVED**

MAR 12 2004

Technology Center 2600

Mail Stop Appeal Brief-Patents  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA. 22313-1450

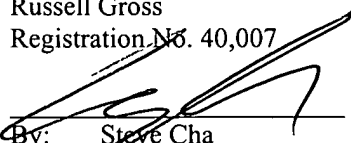
Dear Sir:

Appellants respectfully submit three copies of a Brief For Appellants that includes an Appendix with the pending claims. The Appeal Brief is now due on April 11, 2004.

Appellants enclose a check in the amount of \$330.00 covering the requisite Government Fee.

Should the Examiner deem that there are any issues which may be best resolved by telephone communication, kindly telephone Applicants undersigned representative at the number listed below.

Respectfully submitted,  
Russell Gross  
Registration No. 40,007

By:   
Steve Cha  
Attorney for Applicant  
Registration No. 44,069

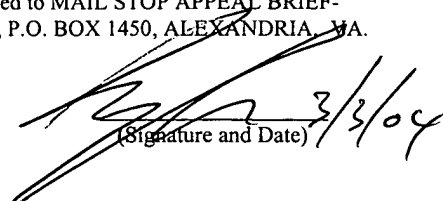
Date: March 3, 2004

**Mail all correspondence to:**  
Russell Gross, Registration No. 40,007  
US PHILIPS CORPORATION  
P.O. Box 3001  
Briarcliff Manor, NY 10510-8001  
Phone: (914) 333-9608  
Fax: (914) 332-0615

Certificate of Mailing Under 37 CFR 1.8

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to MAIL STOP APPEAL BRIEF-PATENTS, COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA. 22313 on March 3, 2004.

Steve Cha, Reg. No. 44,069  
(Name of Registered Rep.)

  
(Signature and Date) 3/3/04



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Before the Board of Patent Appeals and Interferences

In re the Application

Inventor : Mary Benoit

Application No. : 09/494,670

Filed : January 31, 2000

For : DESCRIPTOR FOR A VIDEO SEQUENCE AND  
IMAGE RETRIEVAL SYSTEM USING SAID DESCRIPTOR

RECEIVED

MAR 12 2004

Technology Center 2600

APPEAL BRIEF

On Appeal from Group Art Unit 2613

Date: March 3, 2004

Russell Gross  
Registration No. 40,007  
By: Steve Cha  
Attorney for Applicant  
Registration No. 44,069

03/10/2004 TLWU11 00000041 09494670

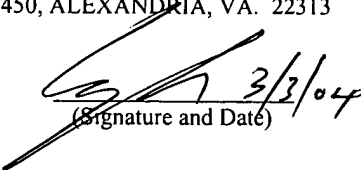
01 FC:1402

330.00 OP

Certificate of Mailing Under 37 CFR 1.8

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to MAIL STOP AF, COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA. 22313 on March 3, 2004.

Steve Cha, Reg. No. 44,069  
(Name of Registered Rep.)

  
(Signature and Date)

## **TABLE OF CONTENTS**

	<b><u>Page</u></b>
<b>I. REAL PARTY IN INTEREST.....</b>	<b>3</b>
<b>II. RELATED APPEALS AND INTERFERENCES.....</b>	<b>3</b>
<b>III. STATUS OF CLAIMS.....</b>	<b>3</b>
<b>IV. STATUS OF AMENDMENTS.....</b>	<b>3</b>
<b>V. SUMMARY OF THE INVENTION.....</b>	<b>4</b>
<b>VI. ISSUES.....</b>	<b>5</b>
<b>VII. GROUPING OF CLAIMS.....</b>	<b>6</b>
<b>VIII. ARGUMENT.....</b>	<b>6</b>
<b>IX. CONCLUSION.....</b>	<b>12</b>
<b>APPENDIX: THE CLAIMS ON APPEAL.....</b>	<b>13</b>

### **I. REAL PARTY IN INTEREST**

The real party in interest is the assignee of the present application, U.S. Philips Corporation, and not the party named in the above caption.

### **II. RELATED APPEALS AND INTERFERENCES**

With regard to identifying by number and filing date all other appeals or interferences known to Appellant which will directly effect or be directly affected by or have a bearing on the Board's decision in this appeal, Appellant is not aware of any such appeals or interferences.

### **III. STATUS OF CLAIMS**

Claims 1-5, 7, 8, 10 and 11 have been presented for examination. All of these claims are pending, stand finally rejected, and form the subject matter of the present appeal.

### **IV. STATUS OF AMENDMENTS**

The Amendment filed September 30, 2003 to which the Final Office Action responds was filed before final rejection of the above-specified claims and has been entered. That Amendment is the second Amendment replying to claim rejections in this application. No subsequent Amendment has been filed.

## **V. SUMMARY OF THE INVENTION**

The selective retrieval of stored image, video and multimedia data is comprehensively, precisely, efficiently and simply accomplished by means of a descriptor in accordance with the present invention (page 2, lines 10-21). The descriptor has a number of formatted fields (page 5, lines 15-18), for whom information is supplied for a particular query on the database to selectively retrieve a desired subset of the video frames in the database (page 1, lines 16-18). Much of the supplied information relates to camera motion in recording the video (page 5, lines 15-18). Camera motion can be fully characterized on the basis of a predetermined number of specific, types of motions (page 1, lines 2-10; FIGs. 1-3). An example of a query for which a descriptor of the present invention is designed is to retrieve a shot, i.e., continuous sequence of video frames (page 5, lines 10-11), that begins with a long zoom of 20 seconds and ends with a short tilt of 2 seconds (page 7, line 12).

Except for “fixed,” each of the motion types can occur in one of two possible directions, e.g., tracking left or right, zooming in or out (page 3, line 33 – page 4, line 3). Each of these bidirectional motion types, according to the instant invention, is therefore oriented and subdivided into two components that stand for two different directions (page 4, lines 3-5). A size of displacement or “magnitude” in each direction assumes a positive number (page 4, lines 3-17). The seven bidirectional motion types yield 14 positive numbers, and a 15<sup>th</sup> positive number represents the “fixed” motion type (page 4, lines 3-5). Since the numbers are each positive, the 15 positive numbers form a histogram representing the magnitude in each direction (page 4, lines 3-5). Magnitude is measured in the number of frames, within a prescribed temporal window of frames, for which the

camera motion occurs in the designated direction (page 4, lines 30-34). The temporal window might begin, for instance, with frame m and conclude with frame m+119, for a total window of 120 frames (page 4, lines 30-34). If, of those frames, only 12 have the camera motion of “zooming in,” then the magnitude in that direction for the prescribed window is 12/120 or 10% (page 4, line 30 – page 5, line 4). A query for which the descriptor temporal presence histogram has a value of 10% in the “zooming in” direction may therefore retrieve frames m through m+119.

Advantageously, in the comprehensive and flexible descriptor, a wide range of temporal granularity is accommodated (page 5, lines 22-24). The temporal window may encompass a whole video sequence, successive shots or even a single frame (page 5, lines 7-14, 22-24). In the case of a single frame, each histogram value is either 0% or 100%, depending on whether the camera motion in the designated direction is present or not (page 5, lines 4-6).

## **VI. ISSUE**

A. Whether claims 1, 5, 7-8 and 10-11 are rendered unpatentable for obviousness within 35 U.S.C. 103(a) over “Video query formulation” by Ahanger et al., “SPIE Proceedings series” (“Ahanger”) in view of U.S. Patent No. 6,389,168 to Altunbasak et al. (“Altunbasak”);

B. Whether claim 2 is rendered unpatentable for obviousness within 35 U.S.C. 103(a) over Ahanger in view of Altunbasak and U.S. Patent 5,267,034 to Miyatake et al. (“Miyatake”); and

C. Whether claims 3-4 are rendered unpatentable for obviousness within 35 U.S.C. 103(a) over Ahanger in view of Altunbasak and U.S. Patent No. 5,929,940 to Jeannin.

## **VII. GROUPING OF CLAIMS**

Claims 1-5, 7, 8, 10 and 11 do not stand or fall together.

## **VIII. ARGUMENT**

Claims 1, 5, 7-8 and 10-11 stand rejected under 35 U.S.C. 103(a) as allegedly unpatentable over “SPIE Proceedings series, 1995” by Ahanger et al. (“Ahanger”) in view of U.S. Patent No. 6,389,168 to Altunbasak et al. (“Altunbasak”).

Claim 1 recites “A video indexing device configured for . . . forming a descriptor that is configured to represent . . . motions of a camera . . . within any sequence of one or more frames of the video scene. . .” The flexibility of the invention for representing a wide range of temporal granularity – this flexibility being evident from the ability to represent even a single frame – is discussed in the specification (e.g., page 5, lines 22-24).

Ahanger, by contrast, fails to disclose or suggest the above-quoted limitation specifically recited in claim 1.

Item 2 of the Final Office Action responds:

Applicant contends that Ahanger fails to disclose “... motion of camera ... within any sequence of one or more frames of the video scene ...” In response, the Examiner respectfully disagrees. Given the claim limitation of one or more frames, the Ahanger reference only has to meet either the one or the more frames. . .

The applicant traverses the latter statement, and, in particular, the latter clause of the latter statement.

Claim 1 recites “A video indexing device configured for . . . forming a descriptor that is configured to represent . . . motions of a camera . . . within any sequence of one or more frames of the video scene . . .”

First of all, Ahanger does not explicitly disclose a descriptor.

Moreover, to the extent, if any, that an Ahanger descriptor can validly be said to be disclosed, such a descriptor could validly be regarded as disclosed only for multiple frames, not for a single frame (Ahanger, section 3, first paragraph, line 6; sixth paragraph (beginning “Another reason . . .”), lines 1-2: “retrieving shots”; seventh paragraph (beginning “By sequential . . .”), line 4). Therefore, in particular, even if Ahanger is deemed to disclose a descriptor, Ahanger makes no disclosure or suggestion of a descriptor configured to represent motions of a camera within a single frame. Ahanger does not disclose that flexibility. Accordingly, the Ahanger “video indexing device” is not “configured for . . . forming a descriptor that is configured to represent . . . motions of a camera . . . within any sequence of one or more frames of the video scene. . .” as explicitly specified in the language of claim 1.

As set forth above, Ahanger fails to disclose or suggest:

A video indexing device configured for receiving a video scene having multiple frames and forming a descriptor that is configured to represent, from a video indexing viewpoint, motions of a camera or any kind of observer or observing device within any sequence of one or more frames of the video scene . . .

as explicitly required by the language of claim 1.

Moreover, Ahanger fails to disclose or suggest:

A video indexing device configured for . . . forming a descriptor that is configured to represent . . . motions . . . comprising at least one of the following basic motion types . . . , wherein each of said motion types, except fixed, is oriented and subdivided into two components that stand for two different directions

Item 3 of the second Office Action, dated June 30, 2003, states that the Ahanger reference shows subdividing in “Figure 1: Basic Camera Operations.”

Figure 1 of Ahanger, however, merely shows that certain camera motion types that have two possible directions. Would an Ahanger “descriptor”, if Ahanger were to even disclose a descriptor, represent zooming out by a positive number and zooming in by a negative number, that number being a single component of a descriptor? We do not know, because Ahanger does not even explicitly disclose a descriptor, even if Ahanger could properly be said to disclose a descriptor at all.

Ahanger makes no disclosure or suggestion whatsoever of a descriptor “configured to represent” “motions” including a motion type that is “oriented and subdivided into two components that stand for two different directions.” “Subdivision” merely exists in the mind of the Examiner using impermissible hindsight gained from reading the present application.

Altunbasak cannot make up for the deficiencies in Ahanger.

Firstly, Altunbasak fails to disclose or suggest “A video indexing device configured for . . . forming a descriptor that is configured to represent . . . motions of a camera . . . within any sequence of one or more frames of the video scene . . .,” and even for this reason alone cannot make up for the deficiencies in Ahanger.

Altunbasak also fails to disclose or suggest that a camera motion type is “subdivided into two components that stand for” two different directions. Accordingly, for at least all of the above-stated reasons, claim 1 is believed not to be rendered obvious by the proposed combination of references.

Claim 10 likewise recites “. . . forming a descriptor that is configured to represent, from a video indexing viewpoint, motions of a camera or any kind of observer or observing device within any sequence of one or more frames of the video scene . . .”

Claim 10 also recites that a camera motion type is “subdivided into two components that stand for” two different directions. Accordingly, claim 10 is believed to distinguish patentably over the applied combination of references for at least the same reasons set forth above with regard to claim 1.

Claim 2 stands rejected under 35 U.S.C. 103(a) as allegedly unpatentable over Ahanger in view of Altunbasak and Miyatake.

As discussed above, Ahanger fails to disclose or suggest “A video indexing device configured for . . . forming a descriptor that is configured to represent . . . motions of a camera . . . within any sequence of one or more frames of the video scene . . .”

Nor does Ahanger disclose or suggest:

A video indexing device configured for . . . forming a descriptor that is configured to represent . . . motions . . . comprising at least one of the following basic motion types . . . , wherein each of said motion types, except fixed, is oriented and subdivided into two components that stand for two different directions

Miyatake operates by correlating the “displacement between frames” (Summary of the Invention: col. 2, lines 37-38), and does not disclose or suggest a descriptor for a single frame. For at least this reason, Miyatake fails to disclose or suggest, alone or in

combination with Ahanger and Altunbasak, “A video indexing device configured for . . . forming a descriptor that is configured to represent . . . motions of a camera . . . within any sequence of one or more frames of the video scene. . .” as explicitly recited in claim 1 and therefore in dependent claim 2. For at least this reason, the proposed combination of references fails to render obvious the invention as recited in claim 2.

Item 5 of the first Office Action states that the Miyatake reference shows subdividing in “Fig. 1, see arrows, e.g.: panning, left or right; and zooming, in or out.”

It is unclear what arrows are being referred to, first because FIG. 1 of Miyatake has no arrows. Presumably, FIG. 3 was intended, but the only arrows in Fig. 3 are those that show the flow of information between components in the Miyatake camera work detection system.

Nevertheless, if, hypothetically, such arrows showing left/right panning or in/out zooming were to exist in a prior art reference, such hypothetical disclosure would merely serve to show that certain camera motion types have two possible directions. In particular, such a showing, if it were to exist, would fall far short of disclosing that a motion type for which a descriptor is configured is “oriented and subdivided into two components that stand for two different directions” as in the present invention.

Referring again to Miyatake, Miyatake averages (col. 9, line 53: “average”) motion vectors to detect a type of camera motion (col. 9, line 51), and therefore does not subdivide the type of camera motion. Miyatake fails to disclose or suggest a camera motion type that is “subdivided into two components that stand for” two different directions.

In addition, Miyatake operates by correlating the “displacement between frames” (Summary of the Invention: col. 2, lines 37-38), and does not disclose or suggest a descriptor for a single frame. For at least this reason, Miyatake fails to disclose or suggest “A descriptor for the representation . . . of camera motions . . . within any sequence of one or more frames” as in the invention recited by claim 1.

For at least all of the above reasons, the proposed combination of prior art references fails to render obvious the invention as recited in claim 2.

Claims 3 and 4 stand rejected under 35 U.S.C. 103(a) as allegedly unpatentable over Ahanger in view of Altunbasak and Jeannin.

Claims 3 and 4 depend from claim 1. Jeannin is directed to estimating motion between segmented images, but cannot make up for the above-described deficiencies in Ahanger and Altunbasak.

Also, for the subject matter particular to claim 4, Official Notice was taken, first in June 30, 2003 Office Action. The replying amendment, of September 30, 2003 traversed the Official Notice. No reference has been asserted to show the purportedly “well-known statement.” In particular, the Final Office Action, dated December 31, 2003 merely reiterates that Official Notice has been asserted and is silent as to any supporting reference.

As to the other rejected claims, each depends from a respective base claim and is deemed to be patentable over the cited prior art at least due to its dependency from its base claim.

IX. CONCLUSION


In view of the above analysis, it is respectfully submitted that the referenced teachings, whether taken individually or in combination, fail to anticipate or render obvious the subject matter of any of the present claims. Therefore, reversal of all outstanding grounds of rejection is respectfully solicited.

Respectfully submitted,

Russell Gross  
Registration No. 40,007

Date: \_\_\_\_\_

3/3/04

  
By: Steve Cha  
Attorney for Applicant  
Registration No. 44,069

## **X. APPENDIX: THE CLAIMS ON APPEAL**

1. A video indexing device configured for receiving a video scene having multiple frames and forming a descriptor that is configured to represent, from a video indexing viewpoint, motions of a camera or any kind of observer or observing device within any sequence of one or more frames of the video scene, said motions comprising at least one of the following basic motion types: fixed, panning (horizontal rotation), tracking (horizontal transverse movement, also called traveling in the film language), tilting (vertical rotation), booming (vertical transverse movement), zooming (changes of the focal length), dollying (translation along the optical axis) and rolling (rotation around the optical axis), or any combination of at least two of these operations, wherein each of said motion types, except fixed, is oriented and subdivided into two components that stand for two different directions, and represented by means of a histogram having a dependent variable with values that each correspond to a respective predefined size of displacement.

2. The device of claim 1, wherein each motion type, assumed to be independent, has its own speed described in an unified way by choosing a common unit to represent it.

3. The device of claim 2, in which each motion type speed is represented by a pixel-displacement value working at the half-pixel accuracy.

4. The device of claim 3, in which, in order to work with integer values, speeds are rounded to the closest half-pixel value and multiplied by 2.

5. The device of Claim 1, wherein a description afforded by said descriptor is hierarchical, by means of a representation of the motion handled at any temporal granularity.

7. An image retrieval system comprising a camera for the acquisition of video sequences, a video indexing device, a database, a graphical user interface for carrying out a requested retrieval from the database, and a video monitor for displaying the retrieved information, an indexing operation within said video indexing device being based on categorization resulting from the use of said descriptor of claim 1.

8. The device of claim 1, wherein the histogram has an independent variable with values configured to each correspond to a different one of said motion types.

10. A computer program product comprising a computer-readable medium having a computer program comprising a sequence of instructions for:

receiving a video scene having multiple frames; and

forming a descriptor that is configured to represent, from a video indexing viewpoint, motions of a camera or any kind of observer or observing device within any sequence of one or more frames of the video scene, said motions comprising at least one of the following basic motion types: fixed, panning (horizontal rotation), tracking (horizontal transverse movement, also called traveling in the film language), tilting (vertical rotation), booming (vertical transverse movement), zooming (changes of the focal length), dollying (translation along the optical axis) and rolling (rotation around the optical axis), or any combination of at least two of these operations, wherein each of said motion types, except fixed, is oriented and subdivided into two components that stand for two different directions, and represented by means of a histogram having a dependent variable with values that each correspond to a respective predefined size of displacement.

11. The product of claim 10, wherein the histogram has an independent variable with values configured to each correspond to a different one of said motion types.